ToxStrategies

Systematic Methods for Assessing Congenital Heart Defects as a Potential Hazard of *in utero* Trichloroethylene Exposures

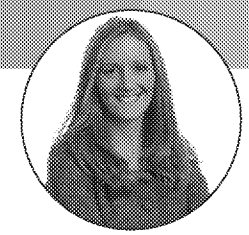
Daniele S. Wikoff[1], Jon D. Urban[2], Laurie C. Haws[2]

[1]ToxStrategies, Asheville, NC;
[2]ToxStrategies, Austin, TX

Innovative Solutions
Sound Science

Daniele Wikoff, PhD

Health Sciences Practice Leader



Dr. Wikoff is a toxicologist and risk assessor with ~15 years professional experience. She has performed evaluation of human health hazards and risks associated with a wide variety of consumer products, food ingredients and additives, pharmaceuticals, and industrial chemicals. Dr. Wikoff is also a practitioner of systematic review and evidence-based methods applied in the fields of toxicology and risk assessment.

- **Practitioner of Systematic Review and Risk Assessment**
  - *Methodology – data quality (risk of bias)*
  - *Design (protocols) and conduct*
  - *Key characteristic of carcinogens*
  - *Endocrine disruption*
  - *Development of health-based benchmarks (e.g., ADI values)*
- **National Academies of Sciences Panel Member and Presenter, Co-Author WHO Guidelines for Systematic Review**
- **Associate Editor – Toxicological Sciences (Systematic Review), Regulatory Toxicology and Pharmacology**
- **Evidence-Based Toxicology Collaboration (EBTC)**
  - *Board of Trustees*
  - *Vice Chair, Science Advisory Committee*
  - *EBTC/EFSA Workshop – Mechanistic Data in Systematic Review (AOPs)*
  - *Study Validity Project Lead*

ToxStrategies

Add stats on pubs, etc?

Jon Urban, PhD, DABT

Health Sciences Assistant Practice Leader

Dr. Jonathan Urban is a board-certified toxicologist > 10 years experience studying and evaluating the potential health effects of a wide range of chemicals of concern (e.g., halogenated aromatic hydrocarbons, solvents, pesticides, metals, hazardous air pollutants), food-related compounds, and consumer product ingredients and contaminants. He is currently supporting the ToxStrategies' efforts developing and applying systematic review methods to the chemical risk assessment process.

Expertise in Toxicology, Risk Assessment and Systematic Review

- Primary lead or contributor on the development and registration of multiple systematic review protocols, as well as peer-reviewed publications
- Studied, refined and applied risk of bias and other data quality tools (e.g., OHAT RoB, TSCA, SciRAP, ToxRTool) in application of chemical risk assessment
- Member of Scientific Review Panel for the National Library of Medicine's Hazardous Substances Databank

Purpose of Presentation

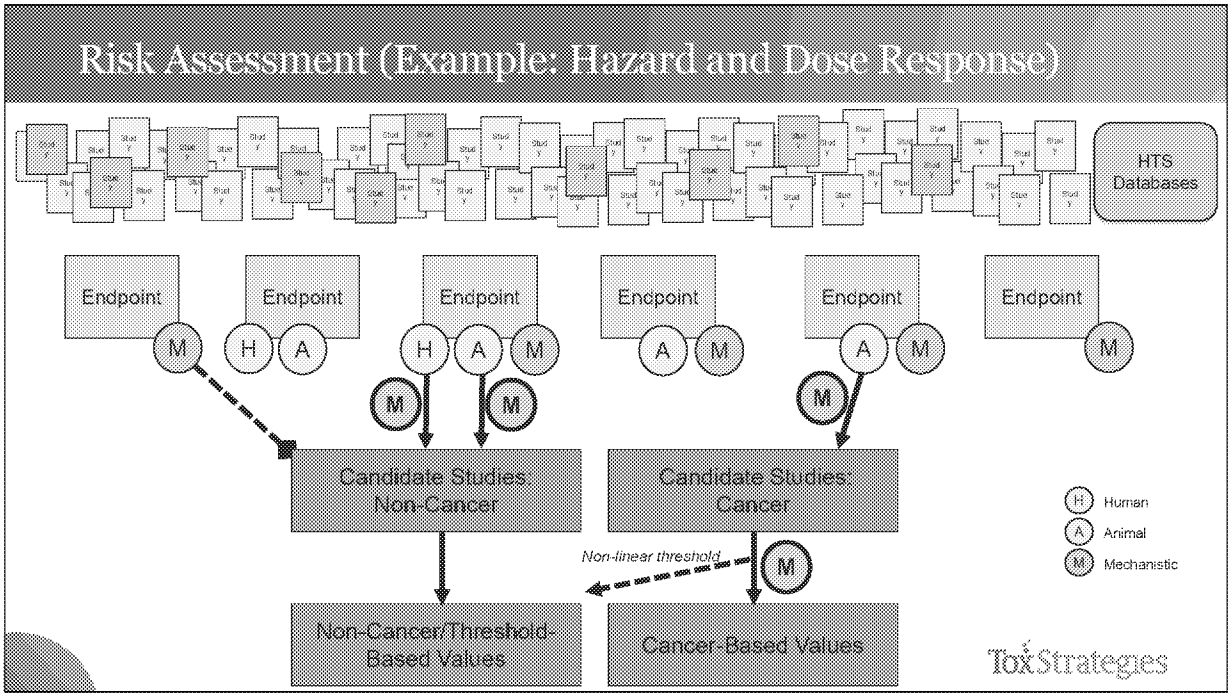
To present findings of a series of exercises carried out by ToxStrategies which were to designed to systematically evaluate whether or not the **overall body of evidence** supports an association between *in utero* TCE exposures increase risk of congenital heart defects (CHDs) in humans

ToxStrategies

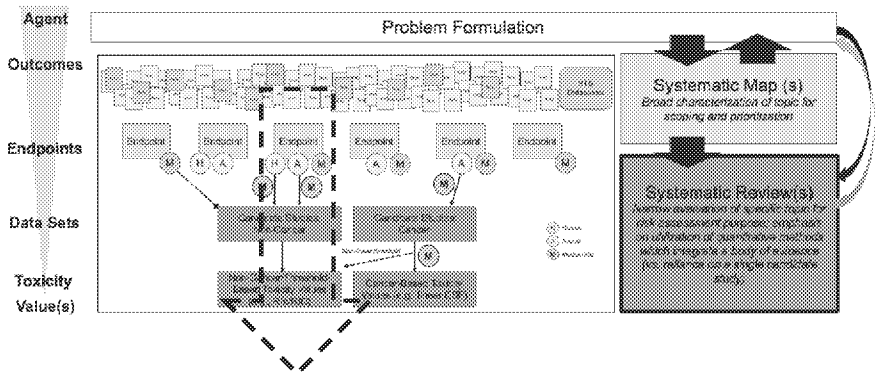
Content Outline

1. Background on systematic review and the application to the TCE-CHD hypothesis
2. Systematic evaluation of human and animal evidence streams relevant to TCE-CHD (Wikoff et al., 2018)
3. Systematic evaluation of mechanistic evidence stream relevant to TCE-CHD (Urban et al., *submitted*)
4. Recent Department of Defense (DoD) systematic review of TCE-CHD evidence base
5. Overall Conclusion
6. Questions

Background on Systematic Review and the Application to the TCE-CHD Hypothesis


Risk Assessment (Example: Hazard and Dose Response)
HTS Databases
Endpoint
Endpoint
Endpoint
Endpoint
Endpoint
Endpoint
M
H
A
H
A
M
A
M
A
M
M
M
M
Candidate Studies: Non-Cancer
Candidate Studies: Cancer
Non-linear threshold
M
Non-Cancer/Threshold-Based Values
Cancer-Based Values
H Human
A Animal
M Mechanistic
ToxStrategies

Evidence-Based Methods Applied to TCE-CHD



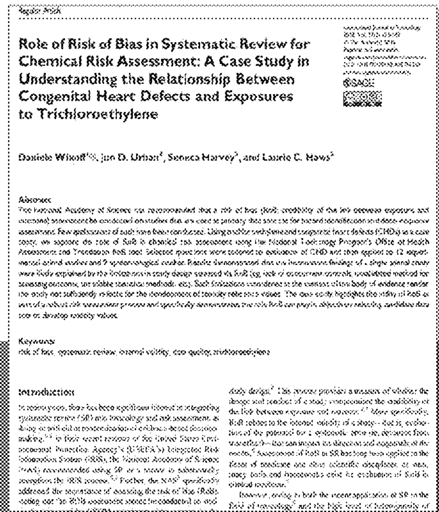
Conducted systematic evaluation of the body of evidence for TCE-CHD by integrating:

1. Human and animal (Wikoff et al. 2018)
2. Mechanistic (Urban et al., submitted)

ToxStrategies

Background: No "Systematic Review" Method Applied to TCE-CHD when Efforts Initiated

	Hardin et al. (2005)	Watson et al. (2006)	Makris et al. (2016)
Review Type:	Narrative	**Weight of the Evidence**	**Weight of the Evidence**
Evidence identification:	No search method described	No search method described	**Select SR elements (e.g., PECO, literature search and screen).**
Evidence reviewed:	Human: occupational and non-occupational, TCE, DCE and general solvents exposures; up to Yauck et al. (2004) and Wennborg et al. (2005). Animal: oral and inhalation studies; up to Johnson et al. (2003) and Zablotny et al. (2002 - SOT poster for later Carney et al., 2006); included Johnson et al. correspondence (2004). Mechanistic: up to Collier et al. (2003) but NOT Ou et al. (2003).	Human: occupational and non-occupational, TCE, DCE and solvent exposures; up to Yauck et al. (2004). Animal: oral and inhalation studies: up to Johnson et al. (2003) and Carney et al. (2001 - ORD lab report for later Carney et al., 2006). Mechanistic: up to Collier et al. (2003) AND Ou et al. (2003).	Human: occupational and non-occupational, TCE exposures: up to Ruckart et al. (2013). Animal: oral and inhalation studies: up to Johnson et al. (2003) and Carney et al. (2006); included Johnson et al. correspondence (2004) and errata (2005, 2014). Mechanistic: up to Makwana et al. (2013).
Study quality evaluation:	No formal criteria applied.	No formal criteria applied.	Criteria developed specifically for review, **informed by 1991 EPA Guidelines for Developmental Toxicity Risk Assessment (non-systematic).**
Mechanistic evaluation:	Evaluated *in ovo*, *in vitro*, *ex vitro* and *in vivo* studies, but not in context of potential TCE-CHD mechanism(s).	Evaluated *in ovo*, *in vitro*, *ex vitro*, and *in vivo* studies [up to Collier et al. (2003) and Ou et al., 2003)]; **assessed in context of six biological processes of cardiodevelopment.**	Evaluated *in ovo*, *in vitro*, *ex vitro*, and *in vivo* studies; **proposed elements of potential adverse outcome pathway (AOP)** based on chick data and mouse KO database
Data integration:	No formal integration approach.	**Bradford Hill's principles of causation.**	**Bradford Hill's principles of causation.**
Conclusion:	"The currently available data do not support TCE and DCE as being specific cardiac teratogens."	"TCE is not a specific cardiac teratogen at environmentally relevant exposures."	"The evidence was characterized as 'Sufficient Experimental Animal Evidence' and 'Limited Human Evidence' [per 1991 EPA Guidelines for Developmental Toxicity Risk Assessment]... the use of the Johnson et al. (2003) study for dose-response assessment remains a reasonable choice."



Role of Risk of Bias in Systematic Review for Chemical Risk Assessment: A Case Study in Understanding the Relationship Between Congenital Heart Defects and Exposures to Trichloroethylene

Daniele Wikoff, Jon D. Urban, Seneca Harvey, and Laurie C. Haws

Systematic evaluation of human and animal evidence streams relevant to TCE-CHD (Wikoff et al., 2018)

ToxStrategies

Objective and Approach

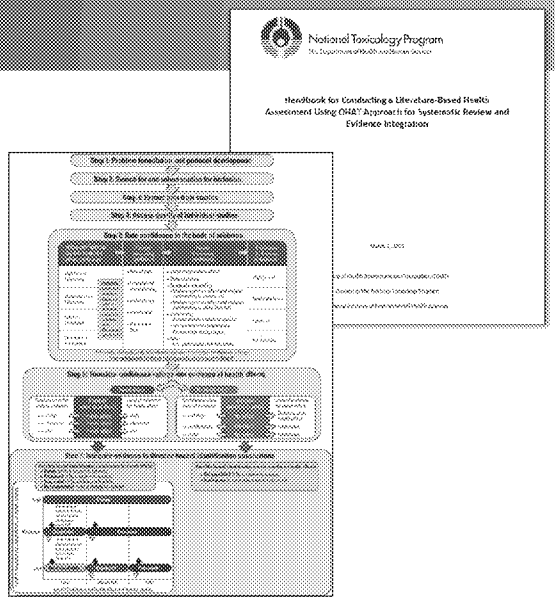
Objective:

To evaluate the internal validity (via Risk of Bias, RoB), as well as other data quality elements, in the human and experimental animal evidence base specific to TCE-CHD and to integrate such into the development of hazard – and risk- based conclusions

Approach:

NTP OHAT Guidance for Systematic Review

- Most complete/final guidance available at the time



PECO (Research Question)

- In humans and experimental animals, is *in utero* exposure to TCE associated with CHDs?
  - Population: human and/or experimental animals
  - Exposure: trichloroethylene via oral ingestion or inhalation
  - Comparator: controls where TCE exposure was absent
  - Outcome: CHDs including defects of the valves (mitral, tricuspid, pulmonary, and aortic), arteries (aorta and pulmonary, including the transposition of major arteries), chambers (atria and ventricular), and septa (atrial, ventricular, and atrioventricular)

Identification of Evidence Base

- Utilized Makris et al. (2016) to identify studies prior to 2015
  - Handsearching of other reviews
- Conducted updated literature search in PubMed & Embase
- Results
  - 12 Experimental animal studies
    (Table 1 in manuscript)
    - 5 oral
    - 7 inhalation
  - 9 Epidemiology studies
    (Table 2 in manuscript)

Experimental Animal Studies (n=12)	
Study Citation	**Basic Study Design Elements**
Oral Studies	
1) Cosby and Dukelow (1992)	Gavage; B6D2F1 mouse (exposed on gestation days [GD] 1-5, GD 6-10, or GD 11-15)
2) Fisher et al. (2001)	Gavage; Sprague-Dawley rat (GD 6-15)
3) Johnson et al. (2003)†	Drinking water; Sprague-Dawley rat (GD 1-22)
4) Narotsky and Kavlock (1995)	Gavage; F344 rat (GD 6-19)
5) Narotsky et al. (1995)	Gavage; F344 rat (GD 6-15)
Inhalation Studies	
6) Carney et al. (2006)	Sprague-Dawley rat (GD 6-20, 6 hr/d)
7) Dorfmueller et al. (1979)	Long Evans rat (GD 1-20, 6 hr/d)
8,9) Hardin et al. (1981)	1) Sprague-Dawley rat (GD 1-19, 7 hr/d) 2) New Zealand White rabbit (GD 1-22, 7 hr/d)
10) Healy et al. (1982)	Wistar rat (GD 8-21, 4 hr/d)
11,12) Schwetz et al. (1975)	1) Sprague-Dawley rat (GD 6-15) 2) Swiss Webster mice (GD 6-15, 7 hr/d)

†Includes Dawson et al. (1993), which first reported the same results from the 2 highest dose groups used by Johnson et al. (2003)

Epidemiology Studies (n=9)	
Study Citation	**Basic Study Design Elements**
Analyses involving direct assessment of TCE and CHD (i.e., offered evaluation and results specific to TCE and CHD)	
1) Bove et al. (1995)§	Cross-sectional, public water system (assumed oral, dermal exposures)
2) Brender et al. (2014)	Case-control, proximity to assumed TCE emitters (assumed inhalation exposures)
3) Forand et al. (2012)	Ecological/Cross-sectional, residential groundwater to vapor intrusion (assumed inhalation exposure)
4) Gilboa et al. (2012)	Case-control, occupational (assumed inhalation and dermal exposures)
5) Tola et al. (1980)	Cohort, occupational (assumed inhalation, dermal exposures)
6) Yauck et al. (2004)	Case-control, proximity to assumed TCE emitters (assumed inhalation exposures)
Analyses involving assessment of exposure to media that may contain TCE	
7) Goldberg et al. (1990)	Case-control (pseudo), residential well (assumed oral, dermal exposures)
8) Lagakos et al. (1986)*	Cross-sectional, residential well (assumed oral, dermal exposures)
9) Ruckart et al. (2013)	Case-control, residential well (assumed oral, dermal exposures)

§Includes Bove (1996), a commonly cited study which reported very similar analyses on the same population
**Later analysis by Massachusetts Department of Public Health not included since complete report was not publicly available*

ToxStrategies

Formal Appraisal of Risk of Bias (RoB) – OHAT with Topic-Specific Refinements

(as recommended by OHAT guidance)

Bias Type	Question No.†	RoB Question	Relevant Study Type
Selection	Q1	Was administered dose or exposure level adequately randomized?	-
		a. Was the the administered dose or exposure adequately randomized among animals?	Animal
		b. Were control and dose groups run concurrently?	Animal
	Q2	Was allocation to study groups adequately concealed?	Animal
	Q3	Did selection of study participants result in appropriate comparison groups?	Human
Confounding	Q4	Did the study design or analysis account for important confounding and modifying variables?	Human
Performance	Q5	Were experimental conditions identical across study groups?	Human
		a. Was the same vehicle used in all study groups?	Animal
		b. Were the non-treatment-related conditions the same for all study groups?	Animal
	Q6	Were the research personnel and human subjects blinded to the study group during the study?	Animal & Human
Attrition/ Exclusion	Q7	Were outcome data complete without attrition or exclusion from analysis?	Animal & Human
Detection	Q8	Can we be confident in the exposure characterization?	Human
		a. Was test article purity reported?	Animal
		b. Was test article solution concentration and stability reported?	Animal
		c. Was test article administered consistently across groups?	Animal
	Q9	Can we be confident in the outcome assessment? - Outcome assessment method reported?	Human
		a. Was the outcome assessment method reported?	Animal
		b. Were the outcome assessors adequately blinded?	Animal
Selective Reporting	Q10	Were all measured outcomes reported?	Animal & Human
Other	Q11	Were appropriate statistical units evaluated and reported?	Animal & Human

++	*Definitely Low* risk of bias: There is direct evidence of low risk of bias practices (May include specific examples of relevant low risk of bias practices)
+	*Probably Low* risk of bias: There is indirect evidence of low risk of bias practices OR it is deemed that deviations from low risk of bias practices for these criteria during the study would not appreciably bias results, including consideration of direction and magnitude of bias
−	*Probably High* risk of bias: There is indirect evidence of high risk of bias practices OR there is insufficient information (e.g., not reported or "NR") provided about relevant risk of bias practices
−−	*Definitely High* risk of bias: There is direct evidence of high risk of bias practices (May include specific examples of relevant high risk of bias practices)

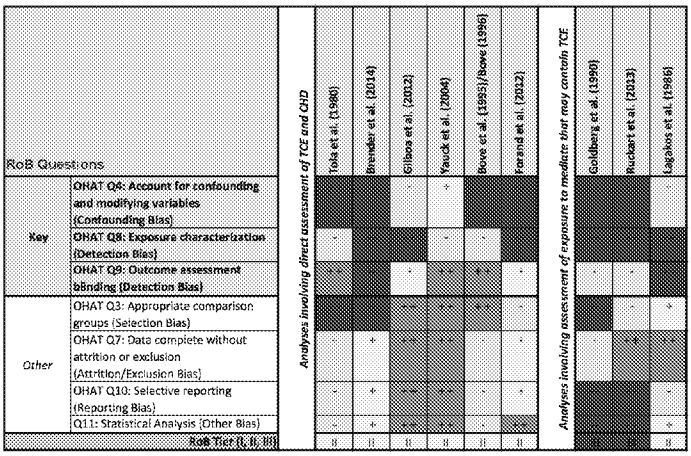
Tier 1: A study must be rated as "definitely low" or "probably low" RoB for key elements and have most other applicable items answered "definitely low" or "probably low" RoB.

Tier 2: A study that neither meets the criteria of tier 1 or tier 3.

Tier 3: A study must be rated as "definitely high" or "probably high" RoB for key elements and have most other applicable items answered "definitely high" or "probably high" RoB.

ToxStrategies

Risk of Bias Evaluation of Human Studies



Studies evaluated in two groups based on directness:

1. Direct evaluation of TCE and CHD
2. No specific evaluation or report of TCE-specific exposures

Key threats to internal validity which limit confidence (*regardless of findings*):

- High RoB in exposure characterization (e.g., use of proximity to TCE sources, residential location, occupational status, etc.)
- High RoB in confounding (e.g., smoking, alcohol, folic acid supplementation)
- Mixed RoB in outcome assessment (e.g., self-reporting)

ToxStrategies

Assessing Confidence in Human Evidence: Very Low to Low

Study Group	Initial Confidence Rating	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Magnitude	Consistency Across Study Types	Final Confidence Rating
Epidemiology studies involving direct assessment of TCE and CHD (i.e., offered evaluation and results specific to TCE and CHD)	Very Low to Moderate	-/↓ Serious (All information is from Tier 2 studies)	-- Inconsistencies assumed to be inherent to study design elements	-- Not serious; direct evaluation of TCE and CHD	-- When provided, confidence interval ranges varied, some were large	-- When effect observed, magnitude was not large	-- Results are not consistent between study types	**Low (++) to Very Low (+) confidence in the human database demonstrating either null or alternative hypothesis**
Epidemiology studies involving assessment of exposure to media with multiple contaminants including TCE (i.e., no evaluation and results specific to TCE)	Very Low to Moderate	↓ Very Serious: Plausible bias that seriously weakens confidence in the results (2 of 3 studies rated Tier 3)	-- Inconsistencies assumed to be inherent to study design elements	↓ Serious: CHD not directly assessed with potential TCE exposure, but more general contaminated water source	↓ Precision cannot be evaluated since no measures of variability were provided	-- Two of three studies reported no effect; when effect observed, magnitude was not large	-- Results not adequately analyzed or reported to evaluate consistency across studies	**Very Low (+) confidence in the human database demonstrating either null or alternative hypothesis**

See Table 2 in Wikoff et al. (2018) for GRADE-based analysis of confidence

Initial confidence ratings ranged from moderate to very low

Confidence further decreased by:

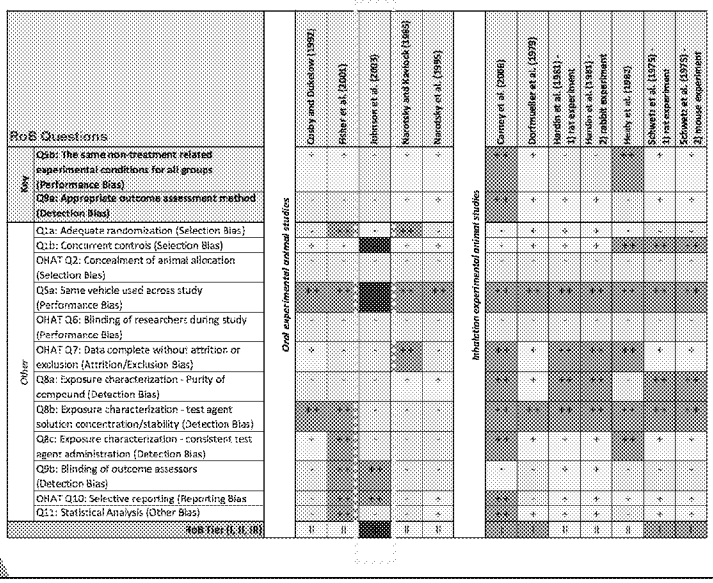
- "serious" and/or "very serious" RoB across evidence base
- inconsistent findings, imprecision, and low magnitude of effects

Resulted in "very low" to "low" level of confidence

Consistent with OHAT methodology, evidence receiving "very low" confidence ratings should not be used to develop conclusions regarding the potential for health effects

Similar conclusion by EPA: "…overall, these epidemiologic studies are not sufficient to establish a causal link between TCE exposure and cardiac defects in humans. This conclusion is consistent with other reviews of the epidemiological literature for TCE exposures and CHD" (Makris et al., 2016)

Risk of Bias Evaluation of Animal Studies



Studies (rat, mouse, and rabbit) evaluated in two groups based on route:

1. Oral
2. Inhalation

Low RoB across most studies; single study identified at Tier III

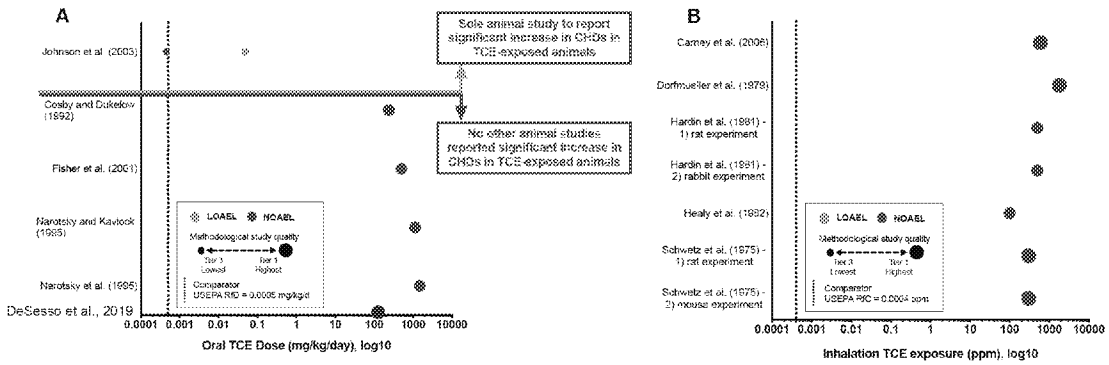
Domains with probably high (or not reported for the element) RoB were similar across studies (i.e., elements of selection and performance bias)

Findings of Tier I and II studies were consistent

- Tier III study was the only one to report positive findings

ToxStrategies

Visual: Validity in Context of Points of Departure



ToxStrategies

NAS Comments on Validity (RoB) and Body of Evidence

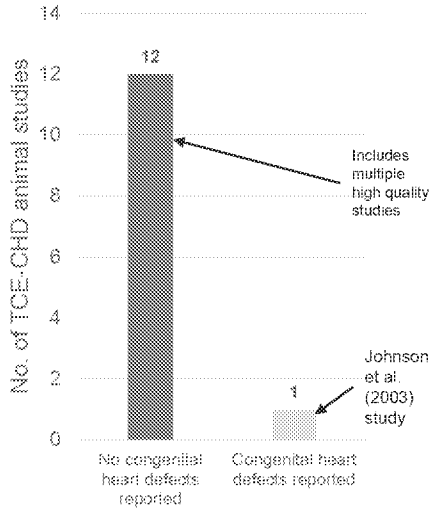
"*The committee is aware that the data on this end point has been controversial, **but found the emphasis on one study to be contrary to systematic review best practices***" – NAS, Comments on DOD OEL (2019)

"*When appropriate, **the committee suggests DOD consider performing dose-response meta-analysis to derive a composite POD for an end point of interest**. A composite POD derived from meta-analyses is based on data from multiple studies, which helps to reduce uncertainty associated with use of a POD from a single study and can increase the overall power to detect an association.*" – NAS, Comments on DOD OEL (2019)

*"Regardless, **a risk of bias assessment should be conducted on studies that are used by EPA as primary data sources for the hazard identification and dose-response assessment**." -- NAS, Review of EPA's IRIS Process (2014)*

"The risk-of-bias assessment can be used to exclude studies from a systematic review or can be incorporated qualitatively or quantitatively into the review results." -- NAS, Review of EPA's IRIS Process (2014)

*".. **the risk-of-bias assessment can be included in the process for selecting studies for calculating toxicity values** or in the uncertainty analysis." – NAS, Review of EPA's IRIS Process (2014)*



ToxStrategies

Can the inconsistent findings of Johnson et al. (2003) be explained?

Results of Wikoff et al. (2018) RoB suggest that inconsistency can be explained by methodological differences:

- High risk of performance, detection, selection, and other (statistical) bias
  - Specifically: lack of concurrent controls, lack of consistent vehicles across control and dose groups, uncertainty in exposures, use of unique and unvalidated outcome assessment method, and pooling of nonconcurrent control group data, etc.

Genetic drift has been proposed as a possible explanation but **is not supported** by the data

- No supporting citations provided by authors that have proposed this as an explanation (Makris et al., 2016; Runyan et al., 2019)
- GLP studies (Fisher et al., 2001; Carney et al., 2006) designed to examine TCE-CHD hypothesis were conducted within a few years of Johnson et al. (2003), not 1-2 decades after
- Incidence of common CHDs (e.g., VSDs) in control Sprague Dawley rats is consistent across multiple breeders located on multiple continents over several decades (DeSesso et al., 2019)
  - **Cardiac development is highly conserved across vertebrate species and unlikely to be affected by genetic drift**

Assessing Confidence in the Animal Evidence: High

Study Group	Initial Confidence Rating	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Magnitude	Dose Response	Consistency Across Study Types	Final Confidence Rating
Oral exposure studies	High	-/↓ Serious (All information is from Tier 2 and 3 studies)	-- Single inconsistency (1/5 oral study) likely explained by study design limitations Inconsistent study could not be validated	-- Not serious; direct evaluation of TCE-CHD	-/↓ The single study to report effect did not report any measure of variability (SD or SE) on CHD findings.	-- No effects observed in 4/5 oral studies; single study reporting effect had low magnitude	-- No TCE-CHD effects observed in 4/5 oral studies; single study reporting effect demonstrated poor dose response	↑ No TCE-CHD effects observed in 4/5 oral studies. Increases confidence in negative findings	**High (++++) confidence in the animal database demonstrating null hypothesis**
Inhalation exposure studies	High	-- Not likely (Most information is from Tier 1 studies)	-- No inconsistency between inhalation studies to explain	-- Not serious; direct evaluation of TCE-CHD	-- No CHDs reported in any of the 7 inhalation studies	-- No effects observed in 7/7 inhalation studies	-- No TCE-CHD response reported in 7/7 inhalation studies	↑ No TCE-CHD response reported in 7/7 inhalation studies. Increases confidence in negative findings	**High (++++) confidence in the animal database demonstrating null hypothesis**

See Table 1 in Wikoff et al. (2018) for GRADE-based analysis of confidence

Initial confidence ratings were high

Confidence not changed or increased

- Low RoB across evidence base
  - Lower RoB in inhalation studies vs. oral studies
- Consistent findings across evidence base
  - Single study reporting effects likely explained by methodological differences

Resulted in "high" level of confidence in findings showing lack of effect

Translation and Integration of Evidence

OHAT Framework: Step 6 - Translate Confidence Ratings into Level of Evidence of Health Effects					TCE-CHD Evidence Base		OHAT Framework: Step 7 - Integrate Evidence to Develop Hazard Identification Conclusions	TCE-CHD Evidence Base	
Confidence in the Body of Evidence		*Direction of effect or no effect*		*Level of Evidence for Health Effect*	*Human Data*	*Animal Data*		Effect/No Effect Level of Evidence by Stream	Overall
(++++) High	➡	Effect/ No effect	➡	High	Low to very low (inadequate) confidence to determine the potential for, or the direction of, an effect (Low/Inadequate)	Very high level of confidence supporting no effect of TCE exposure (oral and inhalation) and CHD	Conclusions for hazard to humans: known, presumed, suspected, not classifiable	Human: Low/inadequate *(Animal: data support no effect)* **= not classifiable**	Not classifiable/not identified to be a CHD hazard to humans
(+++) Moderate	➡		➡	Moderate					
(++) Low	➡		➡	Low			Conclusions for not a hazard to humans: not identified to be a hazard, inadequate to determine hazard	Animal: high Human: Low/inadequate **= not identified/inadequate**	
(+) Very low or no evidence identified	➡		➡	Inadequate					

Conclusions

- Johnson et al. (2003) is a high RoB study and **not reliable for hazard characterization** or for development of noncancer toxicity values. Inconsistent findings of Johnson et al. (2003) can be explained by limitations in study design and reporting
  - High risk of bias, inconsistency, and lack of reproducibility render this study not reliable for hazard characterization and not suitable for selection as a candidate dataset in developing health-based benchmarks

- Human and animal evidence streams **do not support** a relationship between *in utero* exposure to TCE and development of CHDs
  - CHDs are not a suitable end point for risk assessment

Note: Wikoff et al. (2018) analysis does not account for recent OECD Guideline rat drinking water study (DeSesso et al. 2019) that found no evidence of an association between in utero TCE exposure and development of CHDs. However, the results of DeSesso et al. (2019) (no increased CHDs in TCE-treated animals) would not alter Wikoff et al. (2018) conclusions.

Systematic evaluation of mechanistic evidence stream relevant to TCE-CHD (Urban et al., *submitted*)

Background and Objective

- Results from previous assessment (Wikoff et al., 2018) do **not** suggest the need to evaluate the biological plausibility of TCE-CHD as part of a risk assessment
- However, mechanistic literature has been cited as justification for utility of Johnson et al. (2003) and, therefore, ToxStrategies undertook an effort to systematically review the mechanistic evidence stream

Objective: To conduct a systematic evaluation of the mechanistic data related to the TCE-CHD hypothesis, and building on Wikoff et al. (2018), to integrate the synthesis of this evaluation into that of the human and animal evidence streams

Overall Approach For Systematic Evaluation of Mechanistic Data

1. Identify evidence base
2. Evaluate study quality
   - OPPT TSCA SR study quality metrics for *in vitro* and *in vivo* studies
   - Select datasets also assessed using other *in vitro* tools (SciRAP, ToxRTool)
3. Multiple levels of data integration
   1. Within the mechanistic evidence
   2. Combined mechanistic, animal, human
4. Mechanistic evidence conclusions integrated with human and animal evidence streams per NTP-OHAT framework

Characterization of TCE-CHD Mechanistic Evidence Base

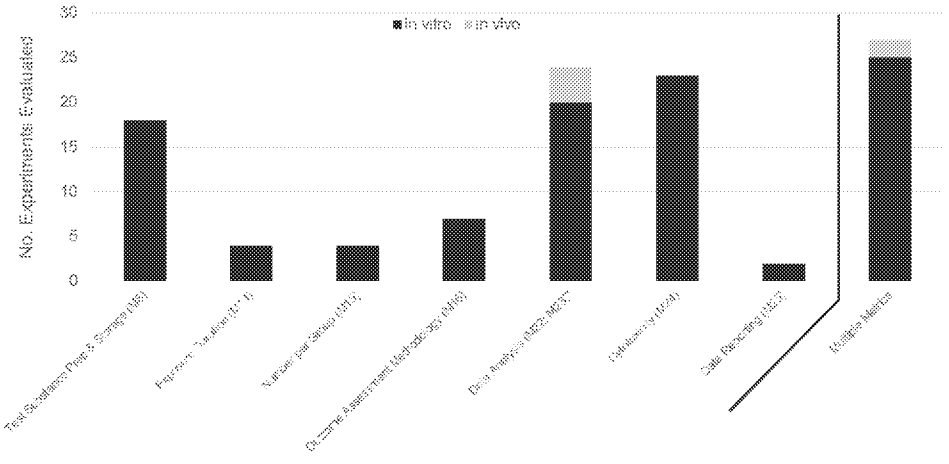
- 22 published studies
- 71 individual experimental datasets
  - 1-7 datasets reported in each study
- Highly heterogenous study designs
  - Many study designs/model: majority (50 of 71) conducted *in ovo* or *in vitro*, but also *ex vivo, ex ovo, in vivo* and recombinant zebrafish experiments
  - Multiple species:
    - In vitro *studies were conducted in cell models from a wide variety of mammalian and non-mammalian species (rat, human, chicken, mouse, bovine, zebrafish)*
  - Endpoints: ranged from molecular (e.g., gene expression, protein interaction), cellular (e.g., morphology, function), organ (e.g., structure, function), and to organism (e.g., viability)
- Findings (e.g., activity, lack of activity) and response directions were varied across the evidence base

Study Quality Evaluation of Mechanistic Evidence Base

Table 1. Summary of data quality metrics for TCE-CHD mechanistic evidence base

Reference	No. of assays relevant to TCE-CHD	≥1 Assay meeting all OPPT study quality metrics?	% TCE-CHD assays meeting all OPPT study quality metrics
Wirbisky et al. (2016)	6 (zebrafish)	Yes (6 of 6)	100%
Drake et al. (2006a)	4 (in ovo)	Yes (4 of 4)	100%
Drake et al. (2006b)	4 (in ovo)	Yes (4 of 4)	100%
Jiang et al. (2016)	4 (in vitro – human)	Yes (4 of 4)	100%
Saillenfait et al. (1995)	3 (ex vivo – rat)	Yes (3 of 3)	100%
Mishima et al. (2006)	2 (ex ovo)	Yes (2 of 2)	100%
Caldwell et al. (2010)	3 (in vivo - mouse)	Yes (1 of 3)	33%
Boyer et al. (2000)	4 (in vitro - chicken)	No	0%
Bross et al. (1983)	2 (in ovo)	No	0%
Caldwell et al. (2008)	2 (in vitro - rat)	No	0%
Collier et al. (2003)	2 (in vivo - rat)	No	0%
Elovaara et al. (1979)	2 (in ovo)	No	0%
Harris et al. (2018)	1 (in vitro – human) 2 (in ovo)	No	0%
Loeber et al. (1988)	1 (in ovo)	No	0%
Makwana et al. (2010)	3 (in ovo)	No	0%
Makwana et al. (2013)	2 (in ovo)	No	0%
Ou et al. (2003)	4 (in vitro – bovine)	No	0%
Palbykin et al. (2011)	3 (in vivo – rat) 4 (in vitro – rat)	No	0%
Rufer et al. (2010)	3 (in ovo)	No	0%
Selmin et al. (2005)	1 (ex vivo – rat) 4 (in vitro – rat)	No	0%
Selmin et al. (2008)	2 (in vitro – mouse)	No	0%
Selmin et al. (2014)	1 (ex ovo) 2 (in vitro – rat)	No	0%
Total (22 References)	*71 Assays*	*7 References, 24 assays*	*34%*

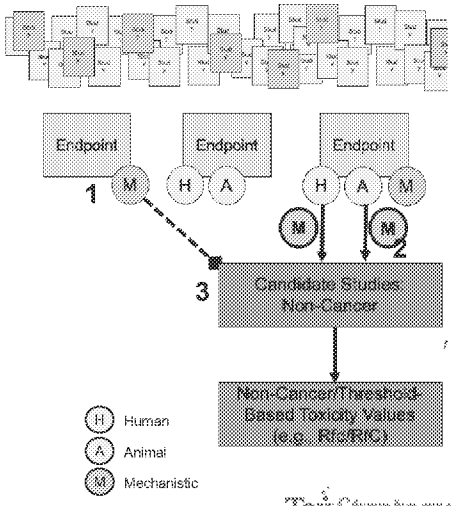
Common Sources of Unreliability Across TCE-CHD Mechanistic Evidence Base (TSCA SR Metrics score = 4)

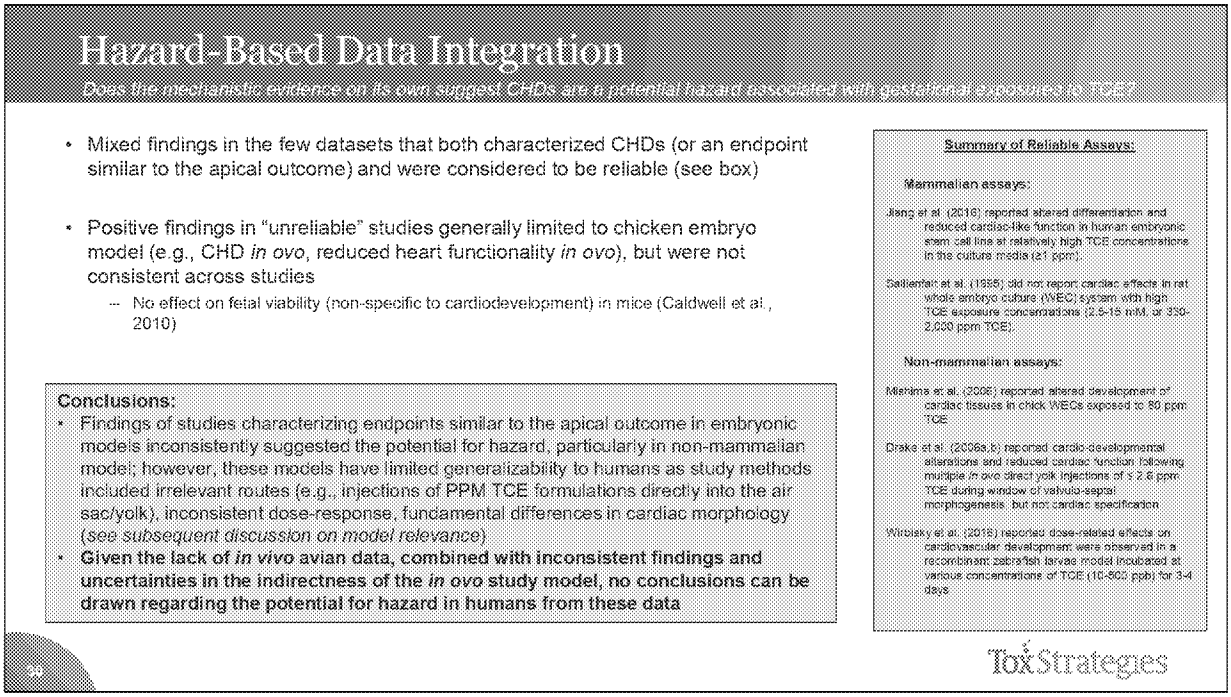


ToxStrategies

Application of Three Evidence Integration Approaches

1. **Hazard-based**: Does the mechanistic evidence on its own suggest CHDs are a potential hazard associated with gestational exposures to TCE?
2. **AOP-based**: Does available mechanistic evidence inform the biological plausibility of TCE-CHD?
   - Mechanistic evidence base insufficient to develop MoA (Makris et al., 2016).
3. **Risk-based**: Do any of the mechanistic studies provide a dose-response datasets that should be considered as candidate studies in developing toxicity values?





Mammalian models limited to human ESCs (reduction in differentiation and cardiac-like functions: Jiang et al., 2006) and rat ex vivo (no cardiac effects in whole embryo cultures: Saillenfait et al., 1995) studies.
Non-mammalian models reported altered cardiac tissue development (in ovo/ex ovo: Drake et al., 2006b; Mishima et al., 2006) or more general altered cardiovascular development (zebrafish: Wirkbisky et al., 2016).

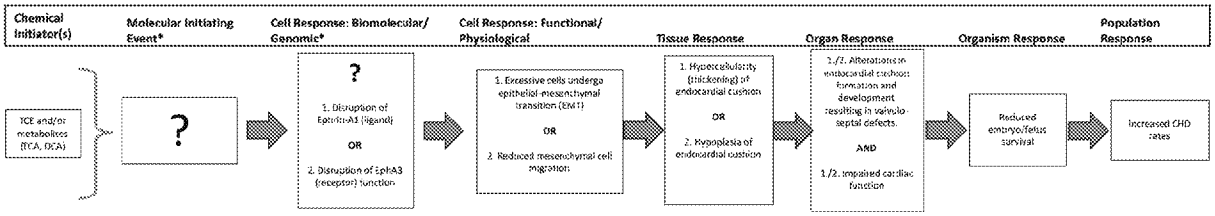
UNRELIABILE STUDIES
In ovo experiments reported increased CHDs (Loeber et al., 1988; Rufer et al., 2010) and reduced heart functionality (Rufer et al., 2010; Harris et al., 2018) following single TCE injections ranging in concentration from 1.3-13 ppm; other in ovo studies (Elovaara et al., 1979; Bross et al., 1983) did not report cardiac defects after injections of considerably higher administered TCE concentrations

Mammalian models More relevant models in mammalian systems did not include observations on cardiac function or structure (Collier et al., 2003; Caldwell et al., 2010; Palbykin et al., 2011), rather were designed to identify changes in fetal heart gene expression.

AOP-Based Data Integration

Does available mechanistic evidence inform the biological plausibility of TCE-CHD?

The putative AOP posited by Makris et al. (2016) was used to organize available mechanistic data, allowing for an evaluation of the biological plausibility of a response in humans:



**Makris et al. (2016) note there is no MEI in this AOP, but speculate on subsequent KE, suggesting that the ephrin-EPH system could "be of high relevance" Currently there are no TCE/TCE metabolite data that indicate a potential MEI or subsequent KE for this theoretical pathway*

AOP-Based Data Integration: Suggestive Evidence

	Molecular Initiating Event (MIE)	Cell Response: Biomolecular/ Genomic	Cell Response: Functional/ Physiological	Tissue Response	Organ Response	Organism Response	Population Response
	Increasing (biological organization)						
Human							
Mouse							
Rat							
Rabbit							
Bovine		Suggestive evidence: Single in vitro assay reports altered protein-protein interactions for one protein proposed in putative AOP.	Suggestive evidence: Single in vitro assay reports reduced cell proliferation.				
Chicken		Suggestive evidence: Single in vitro assay reports reduced EMT biomarker expression. Two in ovo assays report reduced gene and protein expression of potentially relevant marker.	Suggestive evidence: Single in vitro assay reports altered cellular migration, proliferation.	[illegible]	Suggestive evidence in 1 in ovo assay. (5 of 8) reports increased cardiac or altered heart function at low dose. TCE concentrations dependent on timing.	[illegible]	
Zebrafish							

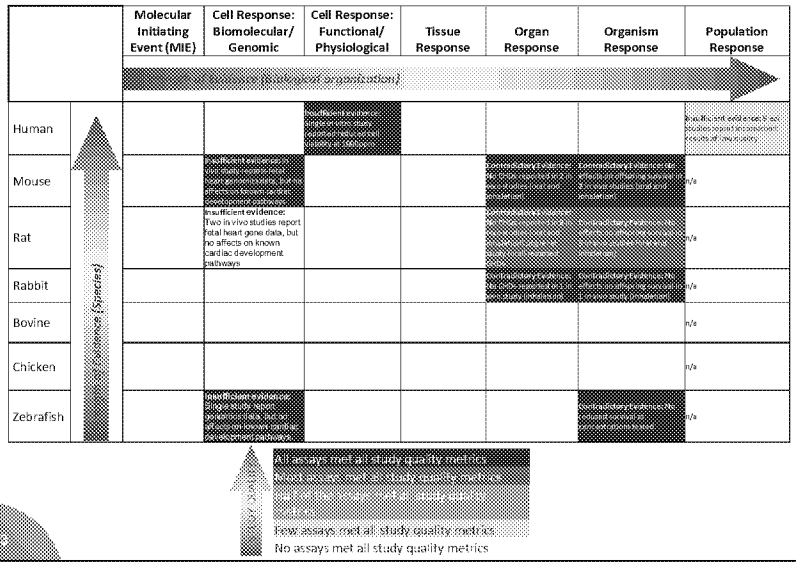
Increasing Confidence (Species)

- All assays met all study quality metrics
- Most assays met all study quality metrics
- Half of the assays met all study quality metrics
- Few assays met all study quality metrics
- No assays met all study quality metrics

- Suggestive evidence limited to experimental assays in bovine cell line and chicken models
- Chicken model provides strongest support of putative AOP for valvulo-septal defects;
  - Chicken most sensitive species tested
- Study quality metrics were consistently met only in studies evaluating responses at the tissue level (altered cardiac cushion cellularity)
  - Lower levels of biological organization less reliable than higher organization levels

ToxStrategies

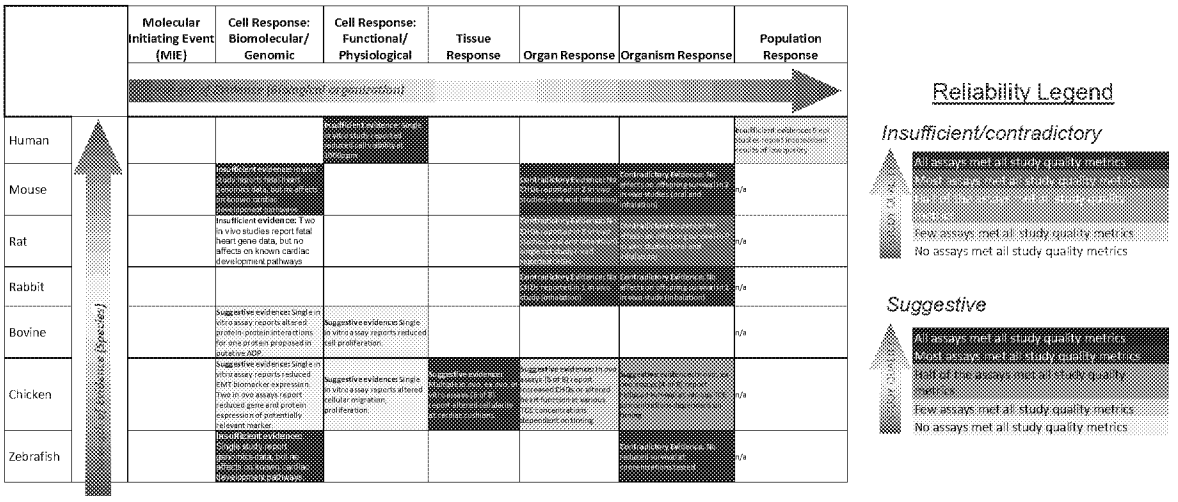
AOP-Based Data Integration: Insufficient or Contradictory Evidence



- Data in humans are insufficient, limited to ESC viability and inconsistent and weak epidemiology studies
- Findings in mammalian models are negative for CHDs *in vivo*, and genomics/gene data are inconsistent and not anchored to adverse effect
- Data in non-mammalian model (zebrafish) limited to genomic and survival data, neither of which support AOP
- Study quality notably higher across these studies than those comprising "suggestive evidence"

ToxStrategies

AOP-Based Data Integration: Combined Findings



Reliability Legend

Insufficient/contradictory

Increasing Quality

- All assays met all study quality metrics
- Most assays met all study quality metrics
- Half of the assays met all study quality metrics
- Few assays met all study quality metrics
- No assays met all study quality metrics

Suggestive

Increasing Quality

- All assays met all study quality metrics
- Most assays met all study quality metrics
- Half of the assays met all study quality metrics
- Few assays met all study quality metrics
- No assays met all study quality metrics

ToxStrategies

AOP-Based Data Integration Observations

Key observations:

1. Approximately half of the available datasets (36 of 71) were considered to be relevant to the putative AOP, of which less than half (17 of 36) met study quality criteria (i.e., were reliable)
2. Chicken embryo appears to be a uniquely sensitive model for TCE-induced CHD, as assays in chicken embryo models (*in ovo, ex ovo, in vitro*) are the only studies which demonstrate activity that would support the plausibility of an effect
   - Majority of evidence supporting the plausibility of TCE-CHD in chickens was considered to be unreliable; only key event that could be considered to have consistent findings based on reliable studies was altered EMT parameters
   - Utility of in ovo model is limited in risk assessment; consistent with other evaluations (e.g., Koustas et al., 2014)
3. Most of the evidence from mammalian models directly contradict the putative AOP
   - Data in humans, mice, rats, and rabbits do not support the plausibility of CHD as a result of *in utero* TCE exposure

Conclusion: The AOP-based synthesis demonstrates a lack of biological plausibility for CHDs associated with TCE exposure in humans (though helps to identify species/model sensitivities in chickens)

Summary of Mechanistic Data Not Related to Putative AOP

(organized by biological level but not by any particular pathway)

	Cell Response: Biomolecular/ Genomic	Cell Response: Functional/ Physiological	Tissue Response	Organ Response
Human	Jiang et al. (2016): In vitro ESC study reports altered effects on expression of cardio-developmental genes (Hlx2.5, Hand1) only at high TCE concentrations (10ppm).	Kang et al. (2015): In vitro "cardiac-like" activities in ESCs (% beating cells; beat rate), reduced ESC differentiation only at high TCE concentrations (21ppm).		
Mouse	Selmin et al. (2008): in vitro PCR 4 genes (not in putative AOP or directly relevant to EMT)			
Rat	Collier et al. (2003): in vivo Serca2a, GPI-p137 gene expression Selmin et al. (2005): in vitro CYP2E1 gene expression (cell & MEC) Selmin et al. (2005): in vitro vimentin gene & protein expression Selmin et al. (2005): in vitro ALDH1 activity Caldwell et al. (2008): in vitro Serca2, Ryr2 gene expression Palbykin et al. (2011): in vitro & in vivo Serca2 protein expression Palbykin et al. (2011): in vitro & in vivo epigenetics (chromatin; SAM availability, Serca2 promoter methylation) Selmin et al. (2014): in vitro protein HNF4a expression Selmin et al. (2014): in vitro HNF4a binding to Serca2 promoter Harris et al. (2018): in vitroHNF4a transcription activity	Caldwell et al. (2008): in vitro Ca2+ homeostasis		
Bovine	Ou et al. (2003): in vitro nitrite, nitrate production Ou et al. (2003): in vitro O2- production			
Chicken	Makwana et al. (2013): in ovo heart CYP2C protein expression Makwana et al. (2013): in ovo heart CYP1A4, CYP2C45, CYP2C gene expression Harris et al. (2018): in ovo heart gene expression (HFN4a, TRAF6, NFKBIE, CYP2C45)	Makwana et al. (2010): in ovo cardiomyocyte contractility		
Zebrafish	Wirbisky et al. (2016): larval PCR 8 genes (not in putative AOP or directly relevant to EMT)		Wirbisky et al. (2016): heart F-actin staining Wirbisky et al. (2016): heart mitochondrial staining	Wirbisky et al. (2016): EGFP imaging of larval vascular system

Note: Experimental data sets that met all reliability criteria in bold

- Majority of "non-AOP" datasets are *in ovo* and *in vitro* models focused on gene or protein expression endpoints
- Rodent and chicken models all published from University of Arizona (UoAZ) labs; investigators hypothesize TCE disruption of Ca2+ homeostasis is key to TCE-CHD hypothesis
- No clear connection between findings of UoAZ studies and remaining *in vitro* models (human ESCs, bovine and zebrafish)

Genomics Data Considerations

Several observations were notable when evaluating genomics studies:

- Frequently cited in support of mechanistic pathways that provide biological plausibility for TCE-CHD (Runyan et al., 2019)
- Most genomics data reported from University of Arizona laboratories
- Reported signals/findings are inconsistent across studies/species (both type and number of altered genes)
- Uncertainty in the strength or specificity of findings: only one of many probes are altered for a gene
- Available data demonstrate that TCE does not alter cardiodevelopmental gateway genes (e.g., gata, nkx2.5, wnt, hand1, etc.)
- Pathway analysis findings do not implicate adverse effects on cardiac development

Conclusions: The inconsistency of the genomics data, combined with the absence of a relevant phenotypic effect ("anchor") in mammalian models (lack of CHDs) significantly limits the utility of these data for informing the potential biological plausibility of TCE-CHDs

Risk-Based Data Integration

Risk-based: *(A set of the mechanistic studies providing a dose-response relationship that should be considered as candidate studies in developing a toxicity value.)*

Note: Risk-based example is **theoretical only, since the overall hazard evidence base shows lack of a relationship between TCE exposure and development of CHDs, and AOP analysis indicates a lack of biologically plausible cardiac defect response in humans.*

Few datasets were of sufficient quality/reliability to consider in risk assessment; only a subset contains sufficient information to characterize the dose/exposure response (include *in ovo*, *in vitro*, and zebrafish models assessing a range of endpoint types - i.e., biomolecular effects at cellular level; cellular morphology, function and proliferation; cardiac function and structure; and organism viability)

Limitations in generalizability/directness (qualitative and quantitative implications)

1. **Dose uncertainty**
   - Ambiguity in the TCE exposures reported in the *in ovo* studies (Drake et al., 2006a,b)
   - Administered TCE concentrations (ranging from 0.13 – 130 ppm TCE) assumed to evenly and instantaneously distribute throughout the egg (one-compartment kinetic model)
2. **Exposure "route" uncertainty**
   - TCE exposure "routes" not physiologically relevant: TCE levels administered *in ovo*, *ex ovo*, and *in vitro* studies cannot account for physiologically relevant *in utero* exposure processes (such as maternal ADME) critical for determining mammalian fetal exposures
3. **Species and biological differences**
   - Developmental timing, gestational physiology, maternal and fetal TCE ADME and toxicokinetics (chicken, zebrafish vs. humans; cell culture vs. *in vivo*)

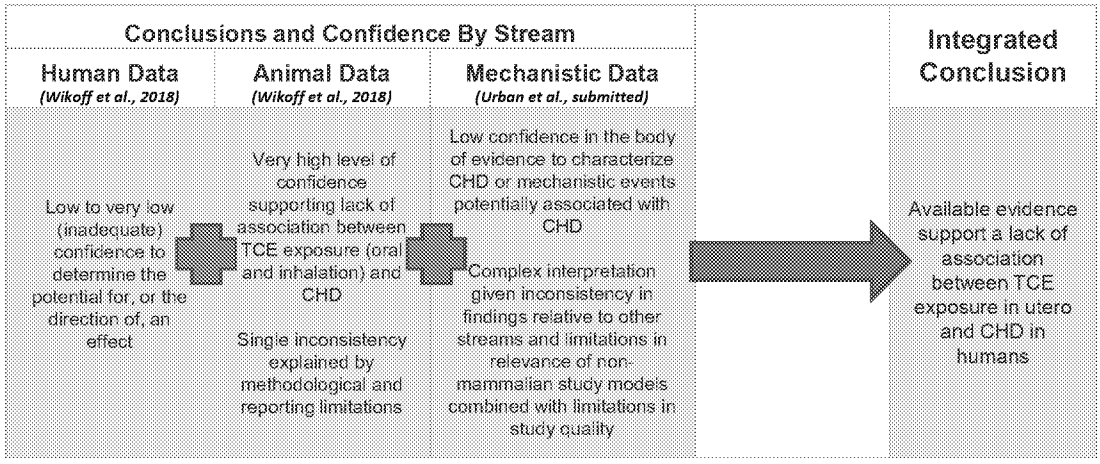
Conclusions:

- Notable absence of empirical knowledge and modeling tools (e.g., *in ovo* to human PBPK, *in vitro*-to-*in vivo* extrapolation) necessary to develop quantitative human exposure estimates with any certainty from mechanistic studies
- The complex challenges and compounding uncertainties that exist in the current TCE-CHD mechanistic dose-response database render these data unsuitable for use in developing toxicity values

Confidence Rating of TCE-CHD Mechanistic Database

Initial Confidence Rating	Risk of Bias	Consistency Across Study Types	Dose Response	Unexplained Inconsistency	Indirectness	Imprecision	Magnitude	Residual Confounding	Final Confidence Rating
Moderate Initial confidence rating based on confidence in exposure (most mechanistic studies fit 3 of 4 categories as "likely"). Most in vitro studies report concentrations at the cell population level, therefore Individual Outcome Data would be similar to human ecological study design for this element (i.e., "may or may not").	↓ 2/3rds of the mechanistic experiments do not meet TSCA study quality standards of "acceptable for risk assessment". Not RoB, but evaluation metrics contain RoB elements, so use as a preliminary measure for this case.	- Results of mechanistic experiments both support and contradict TCE-CHD based on endpoints, though more support than contradict, so leave this as "no impact" (-).	- Some studies that report effects report traditional dose response; however, several also report "hormetic" response (low dose effects that are absent or not as strong at higher doses). Keep as "no impact" (-) since OHAT states this can only increase - not decrease - confidence.	– Inconsistencies likely explained by heterogeneous studies and study design limitations	↓ Nearly half of the mechanistic experiments were conducted in nonmammal models (mostly chicken eggs, one zebrafish), and the "none" of exposure in several mammalian studies (ex vivo whole embryo) were not applicable to human health.	**N/A** Elements of imprecision are accounted for in TSCA study quality metrics (group size, replicates, stats)	– No effects observed in 4/5 oral studies	– Not relevant to animal studies	**Low (++) confidence in the database demonstrating TCE-CHD association: it is likely that in utero TCE exposure is not associated with significant increase in fetal CHD.**

Integration of Mechanistic Evidence with Other Evidence Streams



Overall Conclusions Based on the Body of Evidence

Systematic evaluation of human, animal, and mechanistic streams results in the conclusion that the overall body of evidence does not support an association between TCE exposure in utero and development of CHDs in humans

CHDs are not a suitable endpoint for TCE risk assessment

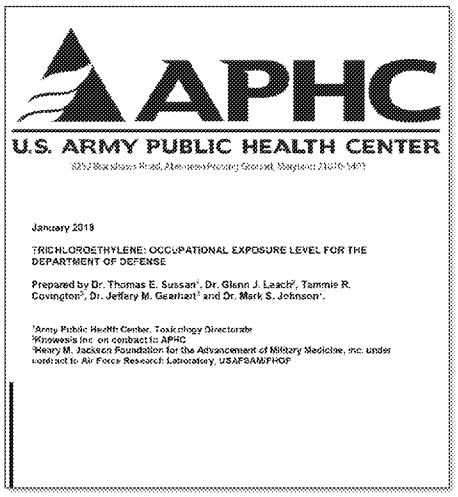
Questions

Recent Department of Defense (DOD) systematic review of TCE-CHD evidence base

2019 TCE Risk Assessment: DoD Occupational Exposure Limit based on Systematic Review of TCE Toxicology Literature

Army Public Health Center (APHC) was tasked with reviewing the TCE non-cancer toxicology and epidemiology literature, to develop an occupational exposure limit (OEL) protective of DoD workers

- Applied systematic review methodology to guide assessment and decision-making process
- APHC developed a tool using elements of other quality assessment tools (e.g., OHAT RoB, ToxRTool, Bradford Hill criteria) to rate the "applicability" of each study (study quality + relevance)
- Resulted in quantitative weigh-of-evidence (WoE) scoring system tailored to fit the purpose and needs of APHC's charge



APHC

U.S. ARMY PUBLIC HEALTH CENTER

5158 Blackhawk Road, Aberdeen Proving Ground, Maryland 21010-5403

January 2019

TRICHLOROETHYLENE: OCCUPATIONAL EXPOSURE LEVEL FOR THE DEPARTMENT OF DEFENSE

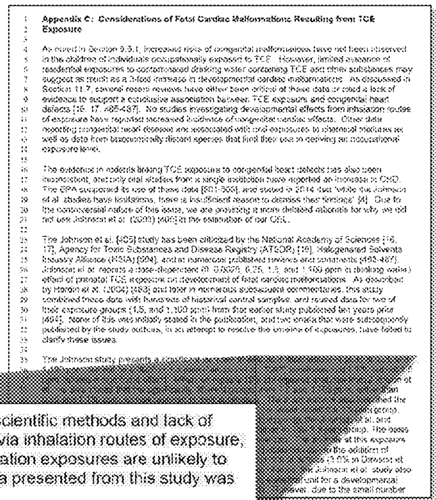
Prepared by Dr. Thomas E. Sussan[1], Dr. Glenn J. Leach[2], Tammie R. Covington[2], Dr. Jeffery M. Gearhart[3] and Dr. Mark S. Johnson[1].

[1]Army Public Health Center, Toxicology Directorate
[2]Wyvesis Inc. on contract to APHC
[3]Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. under contract to Air Force Research Laboratory, USAFSAM/FHOF

ToxStrategies

2019 DoD Systematic Review: Conclusion on TCE-CHDs

- The APHC systematic review included review of developmental endpoints, of which CHDs reported by Johnson et al. (2003) were considered
- APHC summarized several limitations of the Johnson et al. (2003) findings, including additional flaws not previously reported (e.g., dose-response errors, poor BMD fitting)
- They noted that subsequent correspondence and errata (Johnson et al., 2004; 2005; 2014) failed to correct or clarify critical underlying errors



Appendix C: Considerations of Fetal Cardiac Malformations Resulting from TCE Exposure

As noted in Section 9.5.1, increased risks of congenital malformations have not been observed in the children of individuals occupationally exposed to TCE. However, limited evidence of residential exposures to contaminated drinking water containing TCE and other substances may suggest as much as a 3-fold increase in developmental cardiac malformations. As discussed in Section 11.7, several recent reviews have either been critical of these data or cited a lack of evidence to support a conclusive association between TCE exposure and congenital heart defects [16, 17, 485-487]. No studies investigating developmental effects from inhalation routes of exposure have reported increased incidence of congenital cardiac effects. Other data reporting congenital heart disease are associated with oral exposures to chemical mixtures as well as data from taxonomically distant species that limit their use in deriving an occupational exposure level.

The evidence in rodents linking TCE exposure to congenital heart defects has also been inconsistent, and only oral studies from a single institution have reported an increase in CHD. The EPA supported its use of these data [501-503], and stated in 2014 that "while the Johnson et al. studies have limitations, there is insufficient reason to dismiss their findings" [4]. Due to the controversial nature of this issue, we are providing a more detailed rationale for why we did not use Johnson et al. (2003) [405] in the estimation of our OEL.

The Johnson et al. [405] study has been criticized by the National Academy of Sciences [16, 17], Agency for Toxic Substances and Disease Registry (ATSDR) [18], Halogenated Solvents Industry Alliance (HSIA) [504], and in numerous published reviews and comments [485-487]. Johnson et al. reports a dose-dependent (0, 0.0025, 0.25, 1.5, and 1,100 ppm in drinking water) effect of prenatal TCE exposure on development of fetal cardiac malformations. As described by Hardin et al. (2004) [483] and later in numerous subsequent commentaries, this study combined these data with hundreds of historical control samples, and reused data for two of their exposure groups (1.5, and 1,100 ppm) from that earlier study published ten years prior [484]. None of this was initially stated in the publication, and two errata that were subsequently published by the study authors, in an attempt to resolve the timeline of exposures, have failed to clarify these issues.

The Johnson study presents a ...

Conclusion regarding TCE-CHD: The combined influence of these deviations from accepted scientific methods and lack of corroboration with other developmental studies, specifically via inhalation routes of exposure, provide a substantial basis for the conclusion that TCE inhalation exposures are unlikely to cause fetal cardiac malformation in humans. Therefore, data presented from this study was excluded from the quantitative analysis.

ToxStrategies

NAS Review of DOD Assessment

Overall, the NAS was supportive of the DoD's risk assessment conclusions despite making numerous recommendations to refine or improve the systematic review process

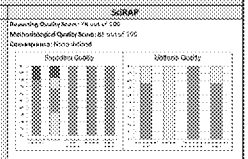
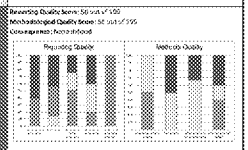
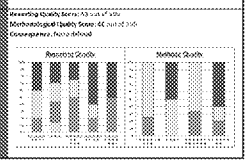
- NAS report echoed many of the written and oral comments submitted by ToxStrategies
  - Apply study quality criteria to epidemiology literature
  - Apply study quality tool to Johnson et al. (2003)
  - Include and assess mechanistic studies in systematic review
  - *Note: ToxStrategies assessments address these critiques*

- Shortcomings in the NAS report:
  - Recommendations that risk of bias be separated from candidate study selection is inconsistent both within the report by the NAS, as well as inconsistent with previous reports from the NAS
  - Critique of DoD study quality criteria are not consistent with previous reports from the NAS, nor are they consistent with criteria being employed by the USEPA (IRIS, TSCA) and other tools (e.g., SciRAP) which recognize elements other than internal validity to be important

Questions

Extra Slides

Similar Trends Across Data Quality Tools (Examples)

Study	Experiment Model & Endpoint	Study Quality Evaluation Outcomes		
		TSCA	TexRTool	SciRAP
Drake et al. (2006b)	**In ovo:** Level of apoptosis in atrioventricular cushion and outflow tract sections of HH24 staged chick embryo hearts	**Score:** 1.3 **Interpretation:** High Quality Study **Consequence:** Retain in TCE CHD evidence base for risk assessment	**Score:** 17 (No "Red" terms) **Interpretation:** Reliable w/o restrictions (RI) **Consequence:** Default check relevance for intended purpose	
Harris et al. (2018)	**In vitro:** Protein activity in transformed human liver cell line (HepG2)	[redacted]	[redacted]	
Sidmon et al. (2005)	**In vitro:** Gene transcription levels in transformed rat heart cell line (H9c2)	[redacted]	[redacted]	

ToxStrategies